

Amélioration de la sécurité chirurgicale avec un jumeau numérique prédictif : le rôle des systèmes multi-agents et de l'apprentissage par renforcement

B. Perez¹

¹ Akkodis Research , AKR. Tour Coeur Défense, 92400 Courbevoie France.

...

bruno.perez@akkodis.com

Résumé

L'article que nous proposons explore les innovations en chirurgie et la nécessité de garantir la sécurité des patients. Malgré les protocoles de sécurité, les risques associés aux interventions chirurgicales deviennent plus complexes. L'utilisation de modèles informatiques prédictifs est suggérée pour améliorer la sécurité. Notre papier aborde notre contribution caractérisée par le couplage d'un système multi-agents (SMA) avec un apprentissage par renforcement pour optimiser la prédiction d'événements indésirables associés aux soins (EIAS). Nous présentons également le concept de jumeau numérique pour simuler et prédire l'état de santé des patients dans un contexte aussi proche de la réalité que possible. Notre article propose ainsi une approche novatrice combinant SMA et apprentissage par renforcement pour créer un simulateur avancé, visant à améliorer la sécurité et l'efficacité des environnements chirurgicaux complexes.

Mots-clés

IA, SMA, AR

Abstract

The article we propose explores innovations in surgery and the necessity of ensuring patient safety. Despite safety protocols, the risks associated with surgical procedures are becoming more complex. The use of predictive computer models is suggested to improve safety. Our paper discusses our contribution characterized by the coupling of a multi-agents system (MAS) with reinforcement learning (RL) to optimize the prediction of adverse events related to care (AERC). We also introduce the concept of a digital twin to simulate and predict the health status of patients in a context as close to reality as possible. Our article thus proposes a novel approach combining MAS and reinforcement learning to create an advanced simulator, aiming to improve the safety and efficiency of complex surgical environments.

Keywords

AI, MAS, RL.

1 Introduction

Dans le contexte actuel des centres hospitaliers, les opérations chirurgicales se déroulent dans un environnement médical qui ne cesse d'évoluer, bénéficiant d'innovations technologiques de pointe et de performances accrues. Cette avancée constante permet non seulement de diversifier les procédures chirurgicales disponibles mais aussi, par voie de conséquence, augmente la complexité des risques associés à ces interventions. La sécurité du patient, primordiale lors de toute intervention chirurgicale, se trouve dès lors intrinsèquement liée aux défis posés par un cadre technique sophistiqué et à la collaboration entre diverses disciplines médicales. L'ontologie médicale, particulièrement développée dans ce secteur, a permis de concevoir des protocoles de sécurité standardisés remarquablement efficaces, déployés entre autres par l'Organisation Mondiale de la Santé (OMS). Malgré l'application rigoureuse de ces processus visant à minimiser les risques, les événements indésirables graves associés aux soins (EIGS) demeurent une réalité préoccupante, avec 2385 cas signalés en 2022 par la Haute Autorité de Santé. Face à ce constat, l'introduction de modèles informatiques prédictifs émerge comme une voie prometteuse pour accroître la sécurité en chirurgie. Dans ce nouveau contexte, l'utilisation d'outils d'aide à la décision apparaît comme une solution d'avenir. Le projet de recherche suggéré ambitionne de développer un jumeau numérique intégrant un système multi-agents et un modèle d'apprentissage continu par renforcement. L'objectif est d'analyser et d'évaluer les risques générés par les différentes actions et conditions au sein du bloc opératoire, afin d'optimiser la prise de décision et de minimiser les risques pour le patient. Nous présentons dans la suite de cet exposé un état de l'art condensé mais représentatif des grandes tendances. Nous décrivons ensuite nos objectifs de recherche en terme de modélisation prédictive dans un contexte connecté avec le réel. La présentation de nos premiers résultats suivi d'une discussion viennent clôturer ce papier.

2 Travaux connexes

Le bloc opératoire est un système complexe où sont pratiqués des actes médicaux susceptibles de produire des effets

indésirables dont la nature et la gravité sont variables. Pour prévenir ces risques, une approche globale est privilégiée, compte tenu de la nature interactive et collaborative du système. Face à cet enjeu de santé publique, de nombreux outils d'évaluation et de prédiction des risques existent et dont les principaux sont présentés dans cet état de l'art. Nous apportons un regard critique par rapport aux convergences et divergences comparativement à notre projet et exposons les verrous soulevés par la littérature.

2.1 Les systèmes multi-agents

Selon Ferber [1] un système multi-agents est un système distribué, constitué d'un ensemble d'agents tels que :

- chaque agent possède des informations et des capacités à résoudre des problèmes limités ;
- le contrôle global du SMA soit impossible ;
- les données soient décentralisées ;
- les calculs soient asynchrones (chaque agent peut effectuer des calculs indépendamment des autres).

L'interaction se caractérise par la mise en relation dynamique de deux ou plusieurs agents qui réalisent un ensemble d'actions réciproques. Elle peut être directe, indirecte, à travers l'environnement, coopérative ou compétitive. La Figure 1 illustre les interactions entre agents, leurs perceptions et leurs actions ; chacun d'eux possédant des objectifs et des connaissances. In fine, un SMA se singularise des autres outils de modélisation en intégrant dans sa dialectique le groupe et donc les notions de coopération, de conflits, de collaboration et de coordination qui en découlent [2]. Les principes fondamentaux des SMA argu-

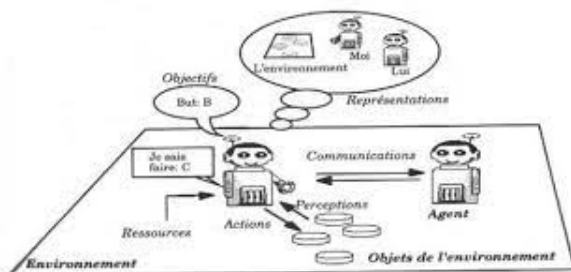


FIGURE 1 – SMA selon Ferber 1995

mentent en faveur d'une modélisation que l'on souhaite non déterministe, à intervalles de temps discret et en capacités d'interactions. Toutefois, il est important de noter une limitation de ces systèmes en ce qui concerne leur capacité à acquérir de nouvelles connaissances. En réalité, ce cadre n'est pas intrinsèquement conçu pour supporter l'apprentissage continu. Dans le paragraphe qui suit, nous explorons plusieurs approches pour intégrer les systèmes multi-agents avec des modèles d'apprentissage.

2.2 SMA et apprentissage

Différents travaux ont été conduits dans le but d'enrichir les capacités d'apprentissage des SMA. Dans son article [3], Marc-Olivier Labarre établit les fondements théoriques de l'apprentissage au sein d'un système multi-agents (SMA),

qu'il divise en deux grandes catégories : les modèles de décision de Markov et les jeux stochastiques. Un modèle de décision de Markov (MDP) génère des issues incertaines suite aux actions de l'agent. Ce modèle est défini par un quadruplet (S, A, R, T) , où S désigne l'ensemble des états possibles, A l'ensemble des actions disponibles, T la matrice de probabilité de transition entre états pour chaque action, et R la matrice de récompenses pour chaque action dans chaque état. À un niveau intuitif, ce processus peut être vu comme un automate qui fonctionne selon des probabilités. L'objectif est de naviguer vers le meilleur état possible, en maximisant les bénéfices et en évitant les états préjudiciables, avec un calcul d'efficacité basé sur les récompenses et les sanctions accumulées au fil des expériences. Nous ne discuterons pas ici des jeux stochastiques, qui ne sont en fait qu'une extension des MDP pour inclure plusieurs agents. Ces stratégies d'apprentissage mettent en lumière une contrainte significative lorsqu'on envisage d'ajouter un élément de prédictibilité. Les mécanismes requis sont en effet plus sophistiqués, nécessitant l'intégration de principes de hiérarchisation et de classification essentiels pour permettre à l'agent d'agir et de prendre des décisions de manière objective. L'apprentissage profond et l'apprentissage par renforcement concurrencent actuellement ces modélisations mathématiques destinées à optimiser le comportement agent. Dans [4], les auteurs proposent un mécanisme unifié permettant aux agents d'un SMA d'apprendre à décider en utilisant l'apprentissage par renforcement profond basé sur une combinaison d'images régénérées contenant des données telles que la vitesse instantanée. En traitant des données de lot séquentielles, les agents sont capables de reconnaître le comportement des agents environnants et de décider de leur propre accélération dans un contexte de circulation routière.

L'ajout d'un jumeau numérique dans ces systèmes, permettant ainsi une connexion avec la réalité physique (jumeau physique), pourrait apporter des améliorations supplémentaires à ces processus, comme il sera expliqué dans le paragraphe suivant.

2.3 Jumeau numérique dans la santé

Les jumeaux numériques sont des supports de simulation utilisés entre autres dans le domaine de la santé pour prédire l'évolution de l'état de santé d'un patient. Ils sont créés en intégrant les données personnelles du patient et en les traitant à l'aide de l'intelligence artificielle. Ces jumeaux numériques sont considérés comme des projections numériques du corps du patient (organe, membre, etc.), fournissant ainsi une représentation virtuelle de son état de santé [5]. L'étude [6] présente une vue d'ensemble des applications actuelles des jumeaux numériques dans les soins de santé, examine les centres de recherche des consortiums et leurs limites, et passe en revue le paysage actuel des nouvelles possibilités de recherche et de développement dans le domaine des soins de santé. De manière générale, les jumeaux numériques appliqués à la santé sont des systèmes complexes peu ouverts sur leur environnement. Cependant, les applications les plus avancées de ces jumeaux numé-

riques se trouvent actuellement dans des domaines très ciblés tel que la cardiologie [7].

Notre étude de la littérature révèle qu'à ce jour, il n'existe pas de jumeau numérique du bloc opératoire. Cette constatation ouvre une perspective que nous avons saisie et qui a justifié notre axe de recherche sur la construction d'un jumeau numérique dont l'architecture est présentée dans la section suivante. Notre objectif est de combler cette lacune dans le domaine et de contribuer à l'avancement de la technologie des jumeaux numériques dans les établissements de santé.

3 Architecture du jumeau numérique

Le couplage SMA/RL, combinant les Systèmes Multi-Agents et l'Apprentissage par Renforcement (AR ou plus communément RL) émerge comme une contribution novatrice à la création de jumeaux numériques, une capacité avancée pour simuler de manière approfondie la complexité des environnements réels. L'intérêt substantiel de cette approche se situe à la fois sur le plan théorique, grâce à la formalisation des interactions entre agents autonomes, et sur le plan pratique, en permettant la simulation avancée de scénarios complexes dans divers secteurs. Sur le plan théorique, le couplage SMA/RL peut être formalisé à travers plusieurs éléments. Considérons les notations suivantes : A : l'ensemble des agents du système. S : l'ensemble des états de l'environnement. R : la fonction de récompense associée aux transitions entre états. P : la dynamique de transition entre les états, déterminée par les actions des agents. Q : la fonction de valeur, représentant l'utilité attendue de chaque action dans un état donné. π : la politique d'action décrivant la stratégie d'un agent. Le couplage SMA/RL peut alors être formellement défini comme la maximisation de la somme des récompenses actualisées sur le temps, sous contraintes de l'équation de Bellman :

Maximiser $\sum_{t=0}^{\infty} \gamma^t \cdot R_t$ sous contraintes de $\mathbb{E}[R_{t+1} + \gamma \cdot \max_{a'} Q(s', a')]$, où \mathbb{E} est l'espérance mathématique, γ le facteur de remise, s un état, a une action, et s' l'état suivant.

Du point de vue pratique, cette formalisation offre des avantages tangibles. En représentant les agents comme des entités dotées d'attributs, de capacités de perception, et de règles comportementales formelles, le couplage SMA/RL permet une simulation avancée. La formalisation des processus d'apprentissage par renforcement, à travers des fonctions de valeur et des politiques d'action, renforce la capacité des jumeaux numériques à modéliser des interactions complexes. Ainsi, cette contribution définie par le couplage SMA/RL constitue un cadre conceptuel robuste pour la modélisation, la simulation, et l'optimisation de notre système complexe. Elle ouvre également la voie à des applications pratiques, notamment dans la robotique, la logistique, la gestion des ressources, en permettant une analyse approfondie des interactions entre agents autonomes dans des environnements dynamiques. La Figure 2 décrit notre architecture dans sa globalité. Le jumeau physique se synchronise avec le jumeau numérique sachant que la communication entre le SMA et le jumeau numérique se fait via une base

pivot qui interagit avec la couche interaction (côté SMA). L'apprentissage par renforcement est quant à lui connecté à la couche d'apprentissage du SMA. Grâce à cet apprentissage, les agents du système peuvent apprendre en continu à partir de l'évolution du système lui-même, ce qui permet de prédire son évolution future et d'optimiser les décisions en temps réel. Cette approche est particulièrement utile dans notre contexte, où les systèmes en question sont en constante évolution et nécessitent une adaptation continue pour fonctionner de manière optimale.

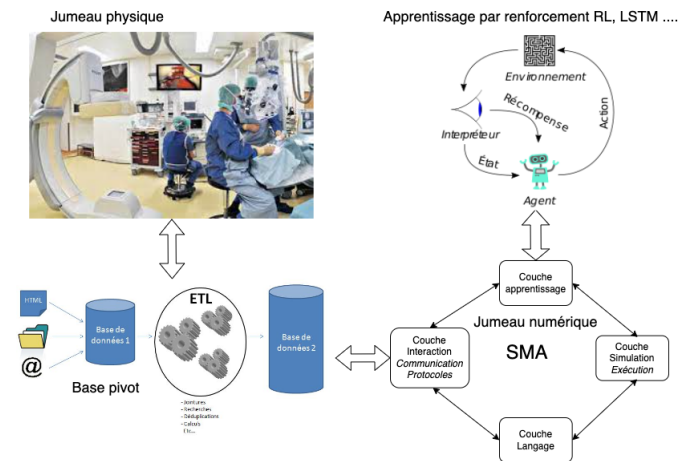


FIGURE 2 – Achitecture globale jumeau numérique

Nous avons porté nos premiers efforts sur la construction du simulateur dans un contexte non connecté et sur l'expérimentation de plusieurs modèles d'apprentissage. Ces travaux sont présentés dans la section qui suit.

4 Nos premiers résultats

L'objet de ce papier ne résidant pas dans une description approfondie de nos premiers travaux nous proposons de décrire respectivement dans les deux paragraphes qui suivent notre première ébauche du simulateur ainsi que nos premières expérimentations de modèles d'apprentissage par renforcement.

4.1 Jumeau numérique : un SMA

Nous nous sommes appuyés sur ces premiers travaux [8] afin de construire notre démonstrateur. Nous avons choisi la plateforme GAMA pour le développement de notre simulateur, car elle offre la possibilité de construire des modèles dans un environnement de développement intégré (IDE) avec le langage GAML (GAMA Modeling Language). Cette caractéristique assure l'enrichissement et l'évolutivité du modèle. De plus, GAMA dispose de plusieurs composants qui permettent d'afficher plusieurs modèles de visualisation dans une même fenêtre. En outre, la plateforme offre la possibilité de construire des modèles très complexes grâce à des outils de gestion de l'espace performants dans différents environnements synchronisés à l'intérieur d'un espace continu de référence. Notre modèle est construit selon l'architecture BDI (Belief-Desire-Intention) : Beliefs représente l'état de connaissance de l'agent sur le monde,

Desires correspond aux objectifs que l'agent souhaite atteindre et Intention sont les action que l'agent décide de mener pour satisfaire ses désirs en fonction de ses croyances. Le tableau 1 donne une vue simplifiée et explicite de cette architecture. "Espèce" correspond aux groupes d'agents (pouvant se limiter à un seul agent) et "Attributs" aux variables d'état des agents. C'est à partir de cette architecture

Espèce	Attributs	Commentaires
Personnel	intention	opérer un patient
	desire	utiliser les ressources
	belief	mesures utiles pour décider
	fatigue	taux de fatigue
	movement	type de mouvement
	qty_mvt	quantité de mouvement
	infected	booléen
Infectious	experience	junior, senior
	type	type agent infectieux
	local	zone d'impact
	intention	contaminer
	desire	utiliser les vecteurs de contamination
	belief	température, pression de l'air
	cont_dist	distance minimum de contamination

Tableau 1 – Tableau récapitulatif architecture BDI.

BDI que nous avons implémenté une première version de notre simulateur. La Figure 3 en donne un aperçu. Le tableau de paramétrage (à gauche de la figure) permet de déterminer le nombre de ressources pour le personnel ainsi que l'environnement infectieux. La couleur rouge associée à un commentaire (sur passage de la souris sur l'un des praticiens) indique si l'agent du groupe "Personnel" est contaminé ou non. Les couleurs vertes, rouges ou bleus de la vue "particles" précisent respectivement si une particule est "non contaminées", "contaminées" ou "décontaminantes". En outre, le simulateur montre l'évolution de la fatigue pour chaque praticien et signale les alertes lorsque les limites sont atteintes. Notre SMA a été enrichie d'un modèle d'apprentissage sélectionné après une comparaison entre plusieurs modèles. Nous détaillerons cette comparaison dans le paragraphe suivant.

4.2 Modèles d'apprentissage par renforcement

Dans le cadre de nos expérimentations, nous avons évalué la performance de notre système multi-agents (SMA) en termes de précision d'apprentissage, de vitesse d'apprentissage, de robustesse. L'objectif étant de comparer les résultats de la simulation du modèle multi-agents avec ceux du jumeau physique simulé numériquement, en utilisant les mêmes entrées et les mêmes conditions initiales. Pour ce faire, nous avons mesuré la précision de la prédiction en utilisant la métrique de l'erreur quadratique moyenne (MSE) entre les résultats de la simulation du modèle multi-agents

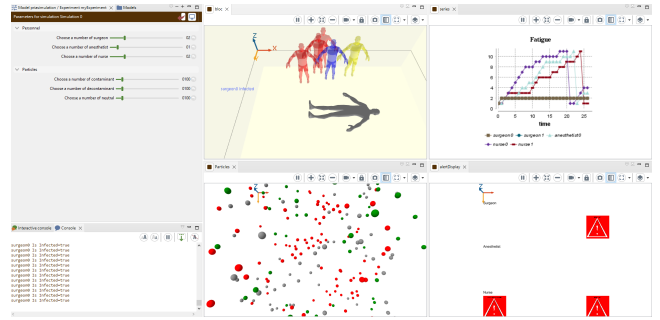


FIGURE 3 – Première version simulateur

et ceux du jumeau physique. Cette évaluation nous a permis de déterminer dans quelle mesure notre modèle multi-agents est capable de prédire avec précision le comportement du jumeau physique, et donc de fournir des informations utiles pour la prise de décision dans un contexte réel. Afin de mener à bien cette analyse comparative, nous avons opté pour l'utilisation de trois modèles d'apprentissage par renforcement : DQN, PPO et A2C. De plus, nous avons également comparé les performances de notre SMA avec un modèle sans apprentissage. Nous avons choisi d'implémenter une modélisation SMA comprenant trois agents : patient (attribut température), praticien (attribut fatigue), espace (attribut taux_contamination). Nous rappelons que les attributs correspondent aux variables d'état. Les algorithmes d'apprentissage par renforcement sont exécutés sur 1000 d'itérations avec un pas de 100 et nos résultats utilisent les métriques suivantes :

- Précision d'apprentissage : récompense moyenne au fil des itérations qui varient dans l'intervalle $[-1, 0]$; les valeurs proches de 0 indiquent un apprentissage optimal.
- Vitesse d'apprentissage : taux de convergence de la récompense moyenne déterminée sur plusieurs points de la courbe de tendance linéaire.
- Robustesse : variation de la précision sur n perturbations (changement de paramètres dans les attributs). Dans notre contexte, la variable n est définie comme égale à 5, ce qui représente le seuil de stabilité.

La Figure 4 récapitule nos expérimentations relatives aux modèles DQN, PPO et A2C. L'axe des abscisses représente les itérations avec un pas de 100, tandis que l'axe des ordonnées affiche les valeurs renvoyées par chaque modèle.

L'apprentissage par renforcement exécuté par l'ensemble des modèles permet d'obtenir des valeurs situées dans l'intervalle $[-1, 0]$, ce qui démontre son efficacité. De plus, les valeurs atteignent un plateau où elles se stabilisent entre $-0,440$ et $-0,327$, soit des valeurs relativement proches de la valeur optimale 0.

A contrario, les résultats obtenus sans apprentissage (absentes du graphique pour des raisons de lisibilité) présentent des valeurs comprises entre $-2,287$ et $-4,0757$, quel que soit le modèle testé en parallèle (DQN, PPO, A2C). Ces valeurs excèdent largement l'intervalle de référence at-

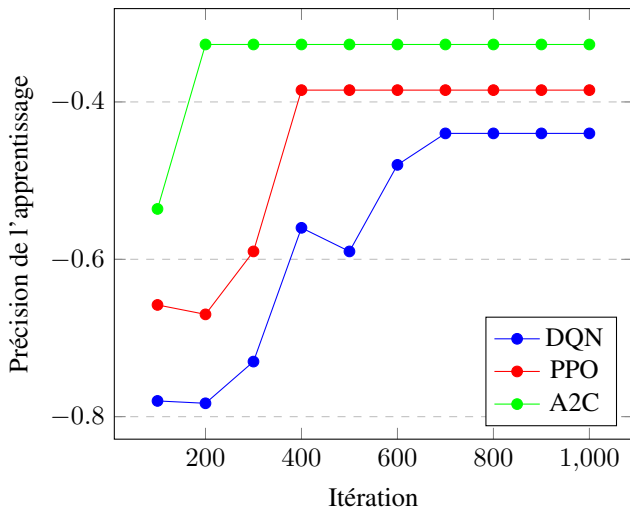


FIGURE 4 – Précision et vitesse d'apprentissage

tendu, à savoir $[-1,0]$.

Comme le montre la synthèse de nos résultats (Tableau 2), le modèle A2C, présente globalement de meilleures performances. Par conséquent, nous avons sélectionné ce modèle pour notre approche d'apprentissage par renforcement opératoire.

Modèle	Précision apprentissage	Vitesse apprentissage	Robustesse
DQN	Très élevée	Lente	Élevée
PPO	Très élevée	Modérée à élevée	Très élevée
A2C	Très élevée	Très élevée	Modérée à élevée

Tableau 2 – Synthèse comparatif modèles apprentissage

5 Conclusion et perspectives

Le projet PRIA vise à créer un système multi-agents (SMA) capable d'apprentissage par renforcement, dans le cadre de notre étude sur la mise en place d'un jumeau numérique d'un bloc opératoire. La modélisation du bloc opératoire s'avère complexe en raison de son contexte dynamique et des interactions directes avec l'évolution du modèle. Pour relever ce défi, nous avons opté pour une modélisation SMA selon l'architecture BDI, qui constitue une première réponse prometteuse. Notre seconde problématique se caractérise par la mise en œuvre d'un apprentissage par renforcement de notre SMA. En effet, la mise en œuvre de ce type d'apprentissage pour notre SMA comporte le risque de converger vers un optimum local plutôt que global, ce qui peut compromettre l'émergence de comportements globaux optimaux. Pour surmonter cette limite, nous avons élaboré une architecture couplant un SMA avec un modèle d'apprentissage par renforcement apte à intégrer des mécanismes de coordination et de communication inter-agents. Cette architecture permet de renforcer la coopération entre

les agents et d'améliorer la prise de décision collective, tout en prenant en compte les contraintes et les incertitudes de l'environnement réel. Nous n'excluons pas dans nos prochaines recherches, d'utiliser des techniques d'apprentissage profond telles que les réseaux de neurones profonds qui permettent de surmonter les limites de complexité et de généralisation des modèles traditionnels d'apprentissage par renforcement.

En terme de perspectives, nous envisageons d'implémenter notre jumeau numérique (SMA "apprenant") avec son jumeau physique qui pour l'instant est simulé. Cela suppose la mise en œuvre de capteurs, de systèmes de synchronisation ainsi qu'une gestion méticuleuse en termes de sécurisation de données.

6 Biblio

Remerciements

Tout d'abord, nous exprimons notre gratitude à l'équipe de recherche pour leur collaboration fructueuse et leur expertise dans les domaines des systèmes multi-agents et de l'apprentissage par renforcement. Nous tenons aussi à souligner l'importance du crédit d'impôt recherche (CIR) dans le financement de ce projet, qui a permis de soutenir les efforts de recherche et de développement dans le domaine de la sécurité chirurgicale.

Références

- [1] J. Ferber, Systèmes multi-agents : Vers une intelligence collective, *Inter Editions*, 1995
- [2] A. Audic, A., and L. Thevin. COrHA : Architecture multi-agents pour la collaboration Utilisateurs-Objets connectés dans une tâche en domotique. *RJCIA'21 (Rencontres Jeunes Chercheurs en Intelligence Artificielle)*, 2021.
- [3] M-O. Labarre, Apprentissage multi-agents. *Département d'informatique et de recherche opérationnelle, Université de Montréal*, <https://www.emse.fr/~boissier/enseignement/sma05/exposes/marcolivier.pdf>, Québec, 2005
- [4] J. Zhong, and H. Hattori, Generation of Traffic Flows in Multi-Agents Traffic Simulation with Agent Behavior Model based on Deep Reinforcement Learning, *arXiv preprint arXiv :2101.03230*, 2020
- [5] Hongchao Shu, Ruixing Liang, Zhaoshuo Li, Twin-S : A Digital Twin for Skull-base Surgery *arXiv :2211.11863 Computer Science, Human-Computer Interaction* 2023.
- [6] Katsoulakis, Evangelia and Wang, Digital twins for health : a scoping review. *Nature Publishing Group UK London* 2024
- [7] J. Tomasi, Impact clinique de la modélisation numérique pour la dissection aortique, *Diss. Université de Rennes*, 2021.
- [8] B. Perez, Prédiction de l'évolution d'un système complexe dans un contexte non déterministe : architecture basée sur le couplage SMA/RàPC et la clusterisation, *Thèse de doctorat. Bourgogne Franche-Comté*, 2021.